# Patient-Reported Outcomes (PROs) In Clinical Trials: Challenges & Opportunities

Nancy Kline Leidy PhD
Sr. Research Leader & Sr VP Scientific Affairs
United BioSource Corporation (UBC)
Bethesda, MD

Anti-Infective Drugs Advisory Committee Meeting
December 9, 2009
Gaithersburg, Maryland

# Overview

- Introductory Comments
  - Scientific Principles - Efficacy; CABP; Health Outcomes, Properties of Study Endpoints
  - Health Outcomes/Endpoints in CABP
  - Properties of Study Endpoints – PROs
  - Clinical Response in CABP
- PRO Instruments & Development/Regulatory Context
  - Development and validation of a PRO
  - Measuring symptoms of CABP – PROs
- Existing PRO Instruments for CABP
  - Pneumonia Symptom Severity Scales (Metlay et al, 1997; Marrie et al., 2004)
  - The Community-Acquired Pneumonia (CAP) questionnaire (el Moussaoui et al., 2004)
  - The Community-Acquired Pneumonia Symptom Questionnaire (Lamping et al., 2002)
  - Next Steps
- Clinical Response in CABP Trials
  - Clinician-Observed Outcomes
  - Standardization – Key Questions

# Scientific Principles

- Efficacy
  - Trials – adequate and well-controlled
  - Design – sample; randomization; masking; non-inferiority or superiority; endpoint positioning
  - Outcome measures – well-defined and reliable
    - » Reliable, valid, responsive
    - » Capture the magnitude of treatment benefit
- CABP
  - Characterized by: selected clinical features (e.g., fever, cough, sputum production & pleuritic chest pain) supported by imaging of the lung, usually chest radiography. Physical examination is supportive (Mandell et al, IDSA/ATS Guidelines 2007)
  - Patients – hospitalized/non-hospitalized
    - » Higher PORT scores = higher risk of mortality
  - Outcomes – success/clinical failure
    - » Mortality (PORT sample enrichment)
    - » ICU Admission
    - » Duration of hospital stay
    - » Clinical Response – signs and symptoms
    - » Resolution – of infectious parameters

# Scientific Principles

- Health Outcomes
  - Results or endpoints of illness – with our without treatment
  - Trial endpoints – with Treatment
    - » Well-defined and reliable
  - Properties of study endpoints
    - » Reliability – Precision
    - » Validity – Measures what it purports to measure
  - Is the endpoint/instrument suitable for a given purpose – e.g., clinical trial?

# Scientific Principles

- Properties of Study Endpoints
  - Reliability – Precision
    - All elements of a given measure correspond/correlate with one another
    - Scores are stable over time in stable patients
    - Scores are reproducible across raters/observers
  - Validity – Measures what it purports to measure
    - Content Validity – Qualitative
      - How well the instrument measures the *target concept*
        - Contains the relevant & important aspects of the concept
        - "What" drives "How"
      - Evaluation – Based on the process used to develop and select items
        - Confidence in the rigor of the development methodology
    - Construct Validity – Quantitative
      - How well scores on the instrument measure (quantify) what is intended
      - Relationship to other outcome measures – similar and dissimilar
        - Known-groups; convergent, discriminant
    - Responsiveness
      - Sensitivity to change

# Health Outcomes/Endpoints in CAPB

- Mortality
  - Sensitivity issue - Small numbers require large samples
  - Validity issue when used alone - Does not assess efficacy outcomes of survivors
- Hospitalization
  - ICU, duration of hospital stay, re-admission rates
  - “Noise” – health policy, hospital policy, clinician practice
- Microbiological response
  - Pathogen eradication
  - “Noise” – Inability to expectorate; no organism identified
  - Validity issue – correspondence to other clinical indicators of resolution
- Chest radiograph response
  - Sensitivity issue – timing
  - Validity issue – correspondence to other clinical indicators of resolution
- Clinical response
  - Time to clinical stability
    - » Vital signs, O2 saturation, IV requirements, mental state
  - Resolution of signs and symptoms
    - » Combination of observed and patient-reported attributes of CAPB
  - Reliability and Validity - ???

# Clinical Response in CAPB Trials

- Signs and Symptoms
  - Sign – objectively observed
    - » Detected by a clinician during a physical examination
  - Symptom – function or feeling experienced by the patient and reported to the clinician
- Clinician Observed:  Signs of pneumonia
  - Fever, increased respiratory rate, increased pulse
  - Low oxygen saturation, cyanosis
  - Decreased breath sounds, bronchial breath sounds, crackles/rales in the upright seated position, egophony
    - » Rarely: vocal fremitus, friction rub, whispered pectoriloquy
  - Dulled percussion over affected lung
  - Variable inter-observer agreement (Metlay et al., 1997; Wipf et al., 1999)
- Patient Reported:  Symptoms of pneumonia
  - Cough, dyspnea, sputum production, pleuritic chest pain, fatigue (Metlay et al., 1997; Marrie et al., 2004)

# Clinical Response in CABP Trials

- Clinical Response:  Patient-reported symptoms and clinician-observed signs of CABP

**Key questions:**

- How is “clinical response” standardized for endpoint measurement?
- How are patient-reported symptoms and clinician-observed signs standardized and quantified to determine “clinical response” to treatment in randomized, controlled trials of CABP treatment in a regulatory context?
- How *should* clinical response be standardized for endpoint measurement in multinational trials?

# Overview

- Introductory Comments
  - Scientific Principles - Efficacy; CABP; Health Outcomes, Properties of Study Endpoints
  - Health Outcomes/Endpoints in CABP
  - Properties of Study Endpoints – PROs
  - Clinical Response in CABP
- PRO Instruments & Development/Regulatory Context
  - Development and validation of a PRO
  - Measuring symptoms of CABP – PROs
- Existing PRO Instruments for CABP
  - Pneumonia Symptom Severity Scales (Metlay et al, 1997; Marrie et al., 2004)
  - The Community-Acquired Pneumonia (CAP) questionnaire (el Moussaoui et al., 2004)
  - The Community-Acquired Pneumonia Symptom Questionnaire (Lamping et al., 2002)
  - Next Steps
- Clinical Response in CABP Trials
  - Clinician-Observed Outcomes
  - Standardization – Key Questions

# Scientific Principles

- Properties of Study Endpoints - PROs
  - Reliability – Precision
    - » All elements of a given measure correspond/correlate with one another
    - » Scores are stable over time in stable patients
    - » Scores are reproducible across raters/observers
  - Validity – Measures what it purports to measure
    - » Content Validity – Qualitative
      - How well the instrument measures the *target concept*
        - Contains the relevant & important aspects of the concept
        - "What" drives "How"
      - Evaluation - Based on the process used to develop and select items
        - Confidence in the rigor of the development methodology
    - » Construct Validity – Quantitative
      - How well scores on the instrument measure (quantify) what is intended
      - Relationship to other outcome measures – similar and dissimilar
        - Known-groups; convergent, discriminant
    - » Responsiveness
      - Sensitivity to change



(Patrick et al., Value in Health 2007, 10: S125-S137)

# The Development & Validation Process

i. Hypothesize Conceptual Framework
- Outline hypothesized concepts & potential claims
- Determine intended application/characteristics
- Develop hypothesized conceptual framework
- Document preliminary instrument development
- Determine the intended population
- Perform literature/expert review
- Position in preliminary endpoint model

ii. Adjust Conceptual Framework & Draft Instrument
- Obtain patient input
- Select recall period, response options & format
- Conduct patient cognitive interviewing
- Document content validity
- Generate new items
- Select mode/method of administration
- Pilot test draft instrument

iii. Confirm Conceptual Framework & Assess Other Measurement Properties
- Confirm conceptual framework with scoring rule
- Finalize instrument content, format, scoring & training
- Assess reliability, validity, sensitivity
- Document measurement development

iv. Collect, Analyze, & Interpret Data
- Prepare protocol & statistical analysis plan
- Evaluate treatment response
- Document interpretation of treatment benefit in relation to claim
- Collect & analyze data



(L. Burke, Measuring Study Endpoints in Clinical Trials, DIA, New Orleans LA 2009)

# The Development & Validation Process: Modified Wheel and Spokes (Simplified)

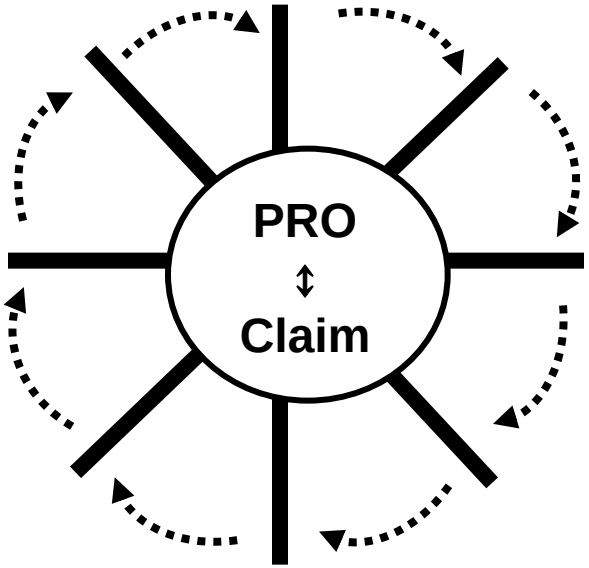
i. Hypothesize Conceptual Framework

ii. Adjust Conceptual Framework & Draft Instrument

iii. Confirm Conceptual Framework & Assess Other Measurement Properties

iv. Collect, Analyze, & Interpret Data

v. Modify Instrument






(L. Burke, Measuring Study Endpoints in Clinical Trials, DIA, New Orleans LA 2009)

# Process and Sample Timelines

- Development 4 – 6 months
  - Literature review
  - Focus groups & interviews
    - » Rate limiting factor – site selection, IRB, recruitment
  - Item pool development
  - Cognitive debriefing
    - » Rate limiting factor – site selection, IRB, recruitment
  - Consultation with experts
- Validation 6 – 18 months
  - Protocol design
  - Study execution
    - » Rate limiting factor – season, site selection, IRB, recruitment
  - Development of the statistical analysis plan
  - Analyses – item reduction and validation
  - Consultation with experts
- Use in clinical trials Ongoing
  - Exploratory or secondary endpoint
  - With experience, use as a secondary or primary endpoint

# Measuring Symptoms of CABP - PROs

- Are there existing CABP Symptom PRO Instruments?
  - Yes
- Can these CABP Symptom PRO Instruments be used in clinical trials evaluating the safety and efficacy of anti-infective agents?
  - Do they follow current standards for endpoint development and validation?
    - » Are the properties consistent with PRO Guidance recommendations?
      Content validity, reliability, construct validity, sensitivity to change?
  - Are they suitable for clinical trials in a regulatory context?
- What are the options?
  - Examine existing instruments for consistency with standards
    - » If consistent, use the instrument
  - Adapt an existing instrument
    - » Make adjustments and validate the modified instrument
  - Develop a new measure
    - » Using current standards and guidance documents

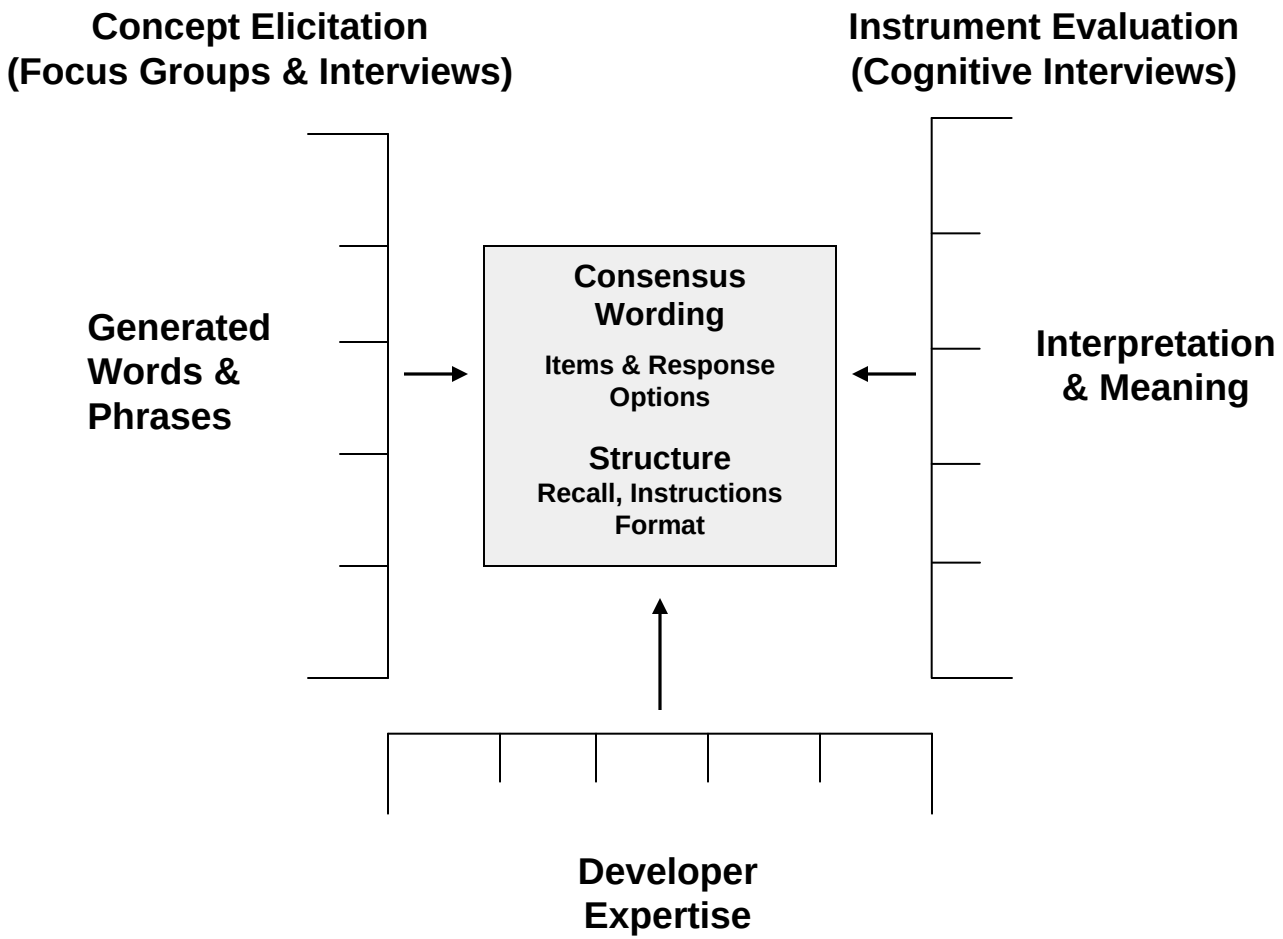
# CAPB Symptom PRO Instruments

- Pneumonia Symptom Severity Scales
  - Symptom Severity Score – Metlay et al., 1997
  - PSS - Marrie et al., J of Infection, 2004
- Community-Acquired Pneumonia (CAP) questionnaire
  - el Moussaoui et al., Thorax, 2004; el Moussaoui et al., Chest, 2006
- Community-Acquired Pneumonia Symptom questionnaire (CAP-Sym)
  - Lamping et al., Chest, 2002; Torres et al., ERJ, 2003

# Properties of CABP Symptom Measures

- Reliability – Precision
  - All elements of a given measure correspond/correlate
  - Scores are stable over time in stable patients
  - Scores are reproducible across raters/observers
- Validity – Measures what it purports to measure
  - Content Validity
    - » The extent to which an instrument contains the relevant & important aspects of the concept it intends to measure.
    - » The items represent a sufficient sampling of content to represent the concept
    - » Evaluation – Based on the process used to develop and select items
      Confidence in the rigor of the methodology
  - Construct Validity
    - » How well the instrument measures what is intended
      Scores represent the outcome
    - » Relationship to other outcome measures – similar and dissimilar
      Concurrent, Convergent, Divergent, Discriminant
  - Responsiveness
    - » Sensitivity to change

# Content Validity:
# Content Consensus through Qualitative Research






ISPOR Task Force on Content Validity of Existing Instruments, 2009 *Value in Health* – Figure 2

# Symptom Assessment in CABP – Content Validity

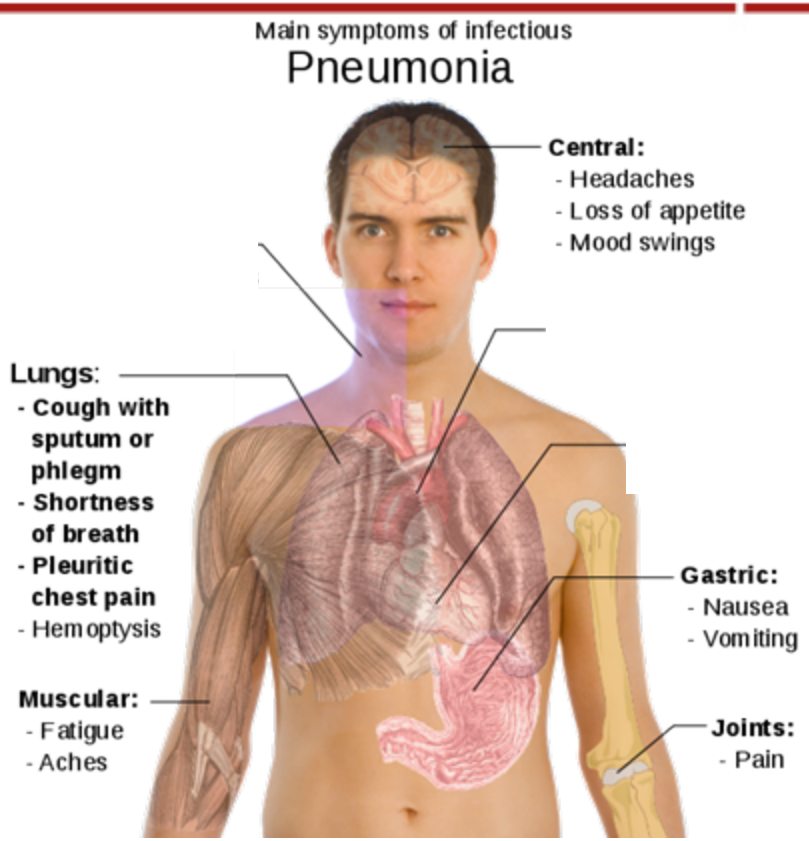
**Literature**

- Cough
- Sputum production (color)
- Dyspnea
- Pleuritic chest pain

Respiratory

- Fatigue
- Tired
- Myalgia/muscle pain
- Headache
- Chills
- Shaking
- Excessive sweating
- Clammy skin
- Nausea
- Vomiting

Systemic






Adapted from: http://en.wikipedia.org/wiki/File:Symptoms_of_pneumonia.svg; public domain

# Symptom Assessment in CABP – Content Validity

**Literature**
- Cough
- Sputum production (color)
- Dyspnea
- Pleuritic chest pain
- Headache
- Chills
- Shaking
- Excessive sweating
- Clammy skin
- Myalgia/muscle pain
- Fatigue
- Tired
- Weak
- Nauea
- Vomiting

**Symptom Severity***

**PSS***

**Symptom Severity****
- ✓ Cough
- ✓ Dyspnea
- ✓ Sputum
- ✓ Pleuritic chest pain
- ✓ Fatigue

* Marrie et al., 2004
** Metlay et al, 1997

**CAP Questionnaire***
- ✓ Cough
- ✓ Sputum production
- ✓ Sputum color
- ✓ Sputum with ease
- ✓ Shortness of breath
- ✓ Severity of shortness of breath
- ✓ Feeling fit
- ✓ General health

* el Moussaoui et al., 2004

**CAP-SYM 12***
- ✓ Coughing
- ✓ Shortness of breath
- ✓ Chest pains
- ✓ Headache
- ✓ Chills
- ✓ Sweating
- ✓ Muscle pain
- ✓ Fatigue
- ✓ Nausea
- ✓ Lack of appetite
- ✓ Trouble concentrating
- ✓ Trouble sleeping

**CAP-SYM 18***
- ✓ Coughing up phlegm
- ✓ Vomiting
- ✓ Coughing up blood
- ✓ Diarrhea
- ✓ Stomach pain
- ✓ Trouble thinking

*Lamping et al., 2002

# Symptom Assessment in CABP – Content Validity

**Literature**
- Cough
- Sputum production (color)
- Dyspnea
- Pleuritic chest pain
- Headache
- Chills
- Shaking
- Excessive sweating
- Clammy skin
- Myalgia/muscle pain
- Fatigue
- Tired
- Weak
- Nauea
- Vomiting

**Symptom Severity***

**PSS***
**Symptom Severity****
- ✓ Cough
- ✓ Dyspnea
- ✓ Sputum
- ✓ Pleuritic chest pain
- ✓ Fatigue

* Marrie et al., 2004
** Metlay et al, 1997

**CAP Questionnaire***
- ✓ Cough
- ✓ Sputum production
- ✓ Sputum color
- ✓ Sputum with ease
- ✓ Shortness of breath
- ✓ Severity of shortness of breath
- ✓ Feeling fit
- ✓ General health

* el Moussaoui et al., 2004

**CAP-SYM 12***
- ✓ Coughing
- ✓ Shortness of breath
- ✓ Chest pains
- ✓ Headache
- ✓ Chills
- ✓ Sweating
- ✓ Muscle pain
- ✓ Fatigue
- ✓ Nausea
- ✓ Lack of appetite
- ✓ Trouble concentrating
- ✓ Trouble sleeping

**CAP-SYM 18***
- ✓ Coughing up phlegm
- ✓ Vomiting
- ✓ Coughing up blood
- ✓ Diarrhea
- ✓ Stomach pain
- ✓ Trouble thinking

*Lamping et al., 2002

# Pneumonia Symptom Score (PSS) (2004)

- Development (content validity)
  - No development history
- Structure
  - 5 Items (fatigue, cough, dyspnea, sputum, pleuritic chest pain)
  - 6-point scale from 0 (no symptom) to 5 (very severe symptom)
  - Patient self-assessment at set intervals, e.g., Days 0 to 14, 30 and 42
- Scoring
  - Sum; range: 0 to 25  (Transformed score 0 to 100)
  - Symptom resolution: total symptom score $\leq$ 20 at day 14 (untransformed score)
    - » "indicates very mild individual symptoms, $\leq$ 1 per symptom)
  - Individual symptom resolution: $\leq$ 1 at day 14
- Context: clinical trial report for resolution of symptoms
  - No data on reliability or validity



(Marrie et al., Journal of Infection 2004; 49: 302-309)

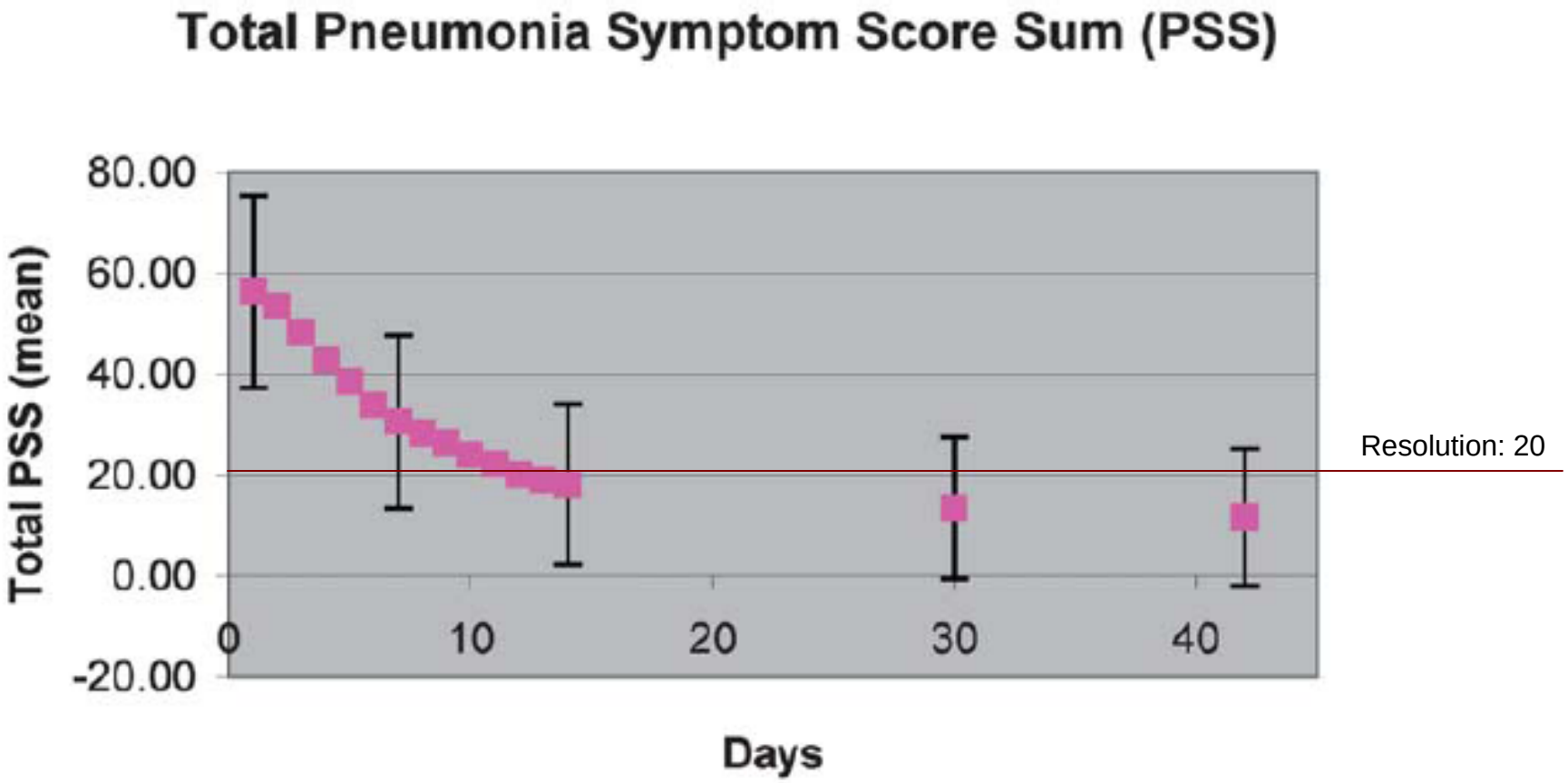
# PSS – Performance Properties

- N=399 CAP
  - Gender: 52% male
  - Mean Age: 48.6 (±15.8) years
- Outpatient
  - Inclusion: Signs and symptoms consistent with mild to moderate bacterial pneumonia not requiring hospitalization, radiologic evidence of new or progressive infiltrate; 2 or more of the following findings: productive cough, purulent sputum, dyspnoea or tachypnea (>20 rr), rigors or chills, pleuritic chest pain
- RCT
  - Efficacy and safety of 2 treatments over 10 days



(Marrie et al., Journal of Infection 2004; 49: 302-309)

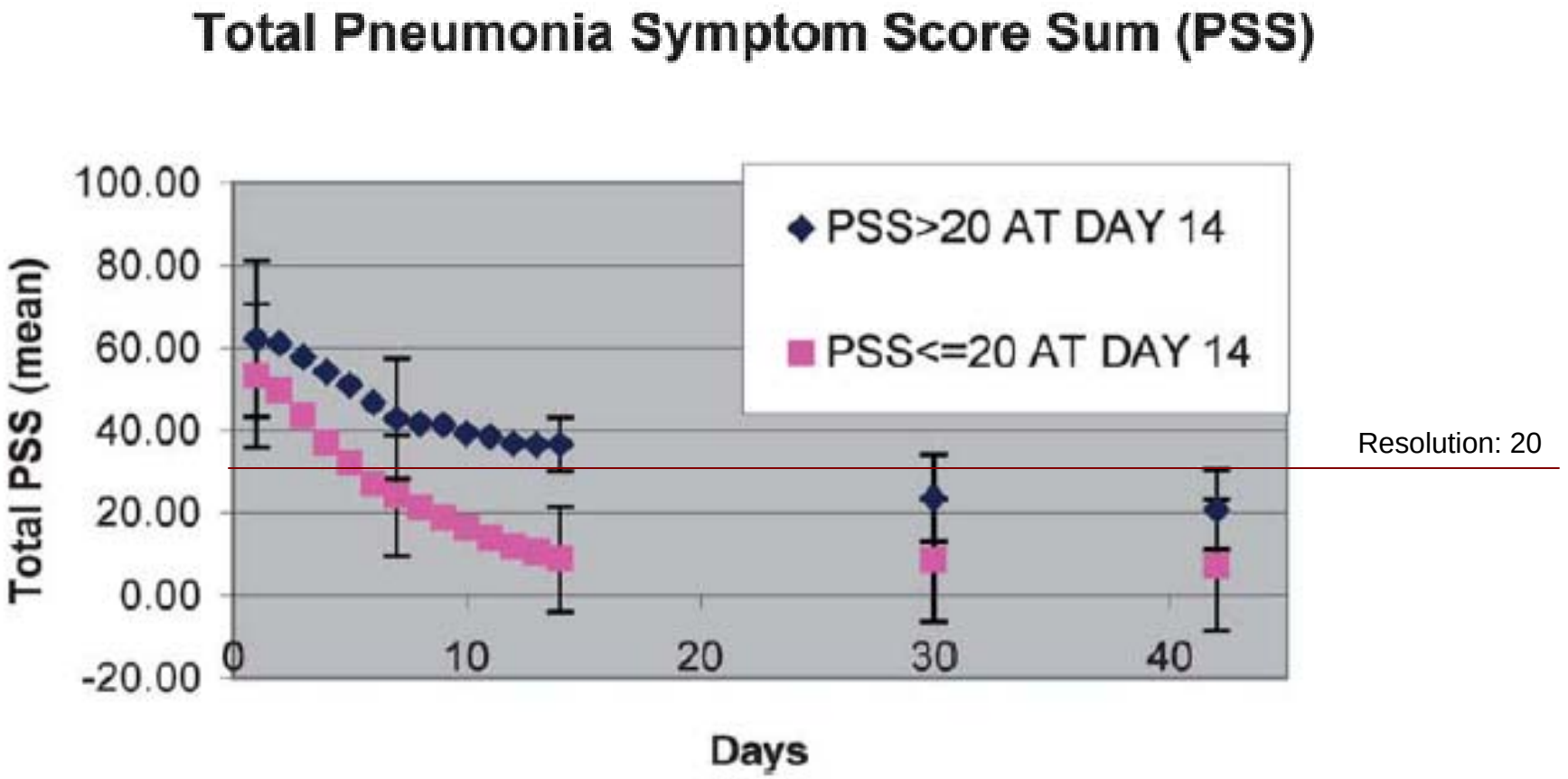
# PSS Change Over Time






(Marrie et al., Journal of Infection 2004; 49: 302-309)

# PSS Change by Resolution Status at Day 14






(Marrie et al., Journal of Infection 2004; 49: 302-309)

# The Development & Validation Process: Modified Wheel and Spokes (Simplified)

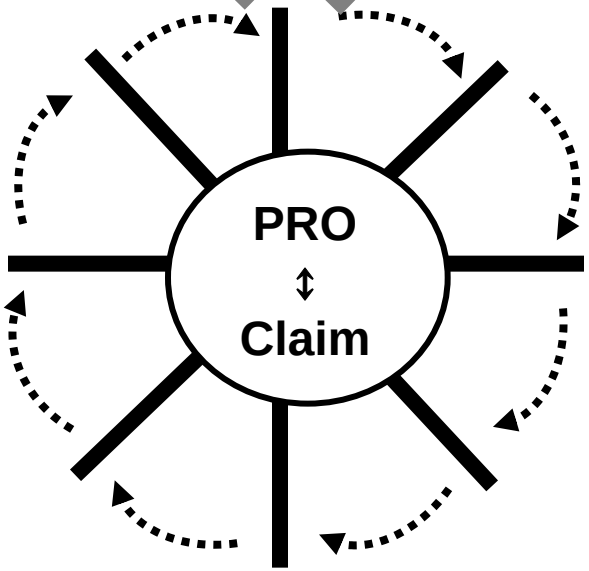
i. Hypothesize C[illegible]eptual Framework

ii. Adj[illegible]ceptual Fram[illegible]k & Draft Instr[illegible]e[illegible]

iii. Conf[illegible]m Conceptual Fram[illegible] & Assess Othe[illegible]urement Prope[illegible]ties

iv. Co[illegible]nalyze, & I[illegible]ret Data

v. Modify [illegible]ment




PSS

# Symptom Severity Score (1997)

- Development (content validity)
  - Panel of investigators; based on prevalent symptoms
  - Response option scaling – based on Anthonisen et al, 1087
- Structure
  - 5 Items (fatigue, cough, dyspnea, sputum, pleuritic chest pain)
  - 2 to 5 point scales; all transformed to 6-point scales (0=none to 5=severe)
  - Mixed mode – interview, mail – at set intervals, e.g., Days 0, 7, 30 90
- Scoring
  - 6-point scale scores summed and transformed to a 0 to 100 summary score
  - Hypothesized meaningful change: 20 points
    - » One symptom change from very severe to absent or all symptoms improving by a single severity point



(Metlay et al., J Gen Intern Med 1997; 12: 423-430)

# Symptom Severity Score – Performance Properties

- N=576 CAP
  - Gender: 38% male
  - Age: 78%<60 years
- Outpatient
  - Inclusion: Acute onset of ≥ 1 of 18 clinical symptoms suggestive of acute illness; radiologic evidence of acute pneumonia within 24 hours o presentation
- Multicenter prospective cohort study
  - Pneumonia Patient Outcomes Research Team (Pneumonia PORT)
- Mode:
  - Mixed interviewer, in-person self, mail survey
  - Days 0 to 7, 30, 90; retrospective recall for pre-pneumonia baseline



(Metlay et al., J Gen Intern Med 1997; 12: 423-430)

# Symptom Severity Score Properties (N=576)

- Reliability
  - Internal Consistency — Cronbach's Alpha: 0.50 (Day 0; 0.70, Day 30 and 90)
  - Test-Re-test — Not reported
- Validity
  - Content Validity — Literature, experts – no patient input
  - Construct Validity
    - » Predictive — Elevated scores at Day 7 or 30 predicted clinic visit
  - Responsiveness — Sensitive to change over time
    Improvement consistent with health status (SF-36)
    Stronger effect size



(Metlay et al., J Gen Intern Med 1997; 12: 423-430)

# Symptom Severity Score – Change Over Time

Table 2A. Proportion Reporting Symptoms During Resolution of Pneumonia*

| | Percentage by Time from Diagnosis | | | | |
|---|---|---|---|---|---|
| Symptom | Prepneumonia | Day 0 | Day 7 | Day 30 | Day 90 |
| Fatigue | 29 | 93 | 80 | 65 | 51 |
| Cough | 16 | 90 | 82 | 53 | 32 |
| Dyspnea | 16 | 68 | 50 | 36 | 28 |
| Sputum | 10 | 63 | 59 | 40 | 27 |
| Pleuritic chest pain | 3 | 47 | 22 | 12 | 8 |

Table 2B. Proportion Reporting Moderate to Severe Symptoms During Resolution of Pneumonia*

| | Percentage by Time from Diagnosis | | | | |
|---|---|---|---|---|---|
| Symptom | Prepneumonia | Day 0 | Day 7 | Day 30 | Day 90 |
| Fatigue | 10 | 79 | 48 | 28 | 20 |
| Cough | 7 | 80 | 51 | 23 | 13 |
| Dyspnea | 2 | 41 | 15 | 7 | 6 |
| Sputum | 3 | 39 | 23 | 12 | 8 |
| Pleuritic chest pain | 1 | 38 | 11 | 5 | 2 |

*n = 576; patients with missing values represented <1% for each time point and were eliminated from those calculations.

Day 0
51.7 (± 20.1)

Day 7
31.2 (± 18.0)

Day 30
19.4 ± 16.9

Day 90
13.6 (± 16.4)



(Metlay et al., J Gen Intern Med 1997; 12: 423-430)

# The Development & Validation Process: Modified Wheel and Spokes (Simplified)

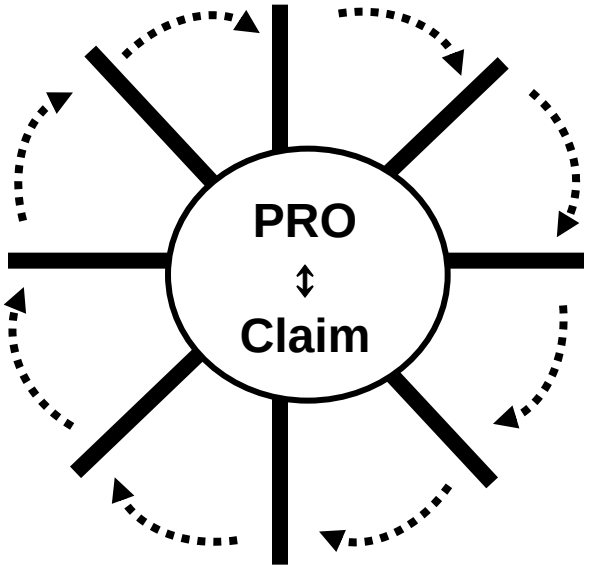
i. Hypothesize Conceptual Framework

ii. Adjust Conceptual Framework & Draft Instrument

iii. Confirm Conceptual Framework & Assess Other Measurement Properties

iv. Collect, Analyze, & Interpret Data

v. Modify Instrument




Symptom Severity

# Symptom Assessment in CABP – Content Validity

**Literature**
- Cough
- Sputum production (color)
- Dyspnea
- Pleuritic chest pain
- Headache
- Chills
- Shaking
- Excessive sweating
- Clammy skin
- Myalgia/muscle pain
- Fatigue
- Tired
- Weak
- Nauea
- Vomiting

**Symptom Severity***

**PSS***

**Symptom Severity****
- ✓ Cough
- ✓ Dyspnea
- ✓ Sputum
- ✓ Pleuritic chest pain
- ✓ Fatigue

* Marrie et al., 2004
** Metlay et al, 1997

**CAP Questionnaire***
- ✓ Cough
- ✓ Sputum production
- ✓ Sputum color
- ✓ Sputum with ease
- ✓ Shortness of breath
- ✓ Severity of shortness of breath
- ✓ Feeling fit
- ✓ General health

* el Moussaoui et al., 2004

**CAP-SYM 12***
- ✓ Coughing
- ✓ Shortness of breath
- ✓ Chest pains
- ✓ Headache
- ✓ Chills
- ✓ Sweating
- ✓ Muscle pain
- ✓ Fatigue
- ✓ Nausea
- ✓ Lack of appetite
- ✓ Trouble concentrating
- ✓ Trouble sleeping

**CAP-SYM 18***
- ✓ Coughing up phlegm
- ✓ Vomiting
- ✓ Coughing up blood
- ✓ Diarrhea
- ✓ Stomach pain
- ✓ Trouble thinking

*Lamping et al., 2002

# CAP Questionnaire

- Development (Content Validity)
  - Textbooks, literature, experts
  - “The most specific symptoms that characterise the respiratory condition in CAP”
- Structure
  - 9 items
  - Scaling:
    - » Dyspnea – yes/no
    - » Fatigue and fitness – VAS
    - » Others – Likert-type scale (ordinal scaling)
  - Scoring
    - » Total score; respiratory score; well-being score



(el Moussaoui et al., Thorax 2004; 59:591-595)

# CAP Questionnaire – Performance Properties

- N=67 CAP
  - Gender: 67% male
  - Mean Age: 56 (17.8) years  Range 21-96
  - PSI 56 (23.4) – Range: 20-106
- 4 of 8 study hospitals
  - Inclusion criteria: temp >38; clinical signs of pneumonia, new infiltrate on chest radiograph, PSI < 110
- RCT
  - Comparing 2 durations of treatment of CAP



(el Moussaoui et al., Thorax 2004; 59:591-595)

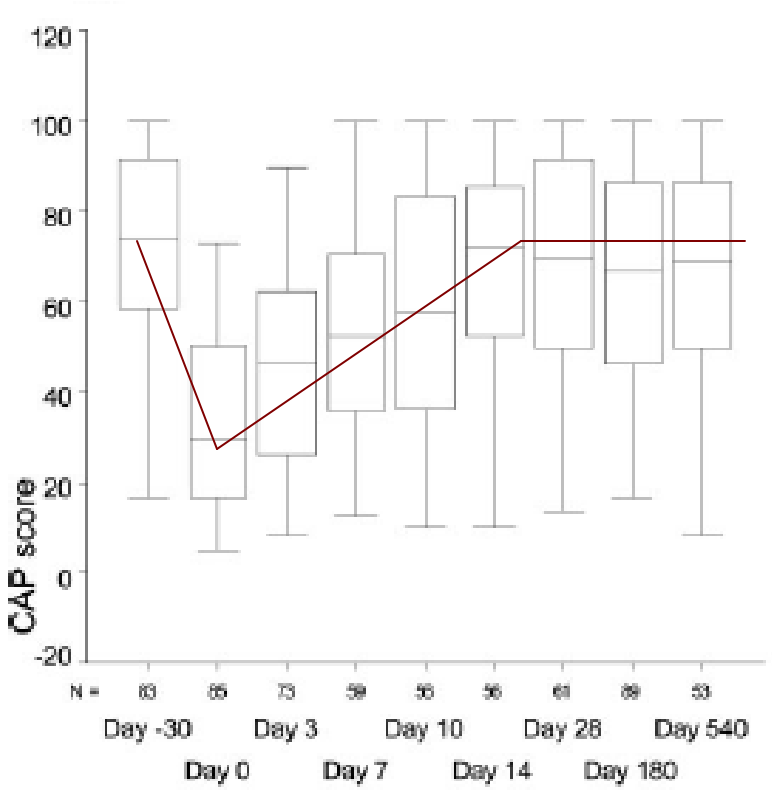
# CAP Questionnaire Performance Properties (N=67)

- Reliability
  - Internal Consistency — Cronbach's Alpha: 0.87
  - Test-Re-test — ICC: 0.83
- Validity
  - Content Validity — Literature, experts – no patient input
  - Construct Validity
    - » Within Scale Analyses — Alpha=0.87
    - » Clinical — Corr with physician judgment (r=0.35), temp (r=-0.43); respiratory rate (r=-0.34), O2 sat (r= 0.23) WBC (r=-0.25); CRP (r=- 0.31); ESR (r=-0.17)
  - Responsiveness — Change from Normal to Baseline; Baseline to day 10; baseline to day 28 (ES $\geq$ 1)



(el Moussaoui et al., Thorax 2004; 59:591-595)

# CAP Questionnaire Change Over Time




(el Moussaoui R et al., Chest, 2006; 130: 1165-1172)

# The Development & Validation Process: Modified Wheel and Spokes (Simplified)

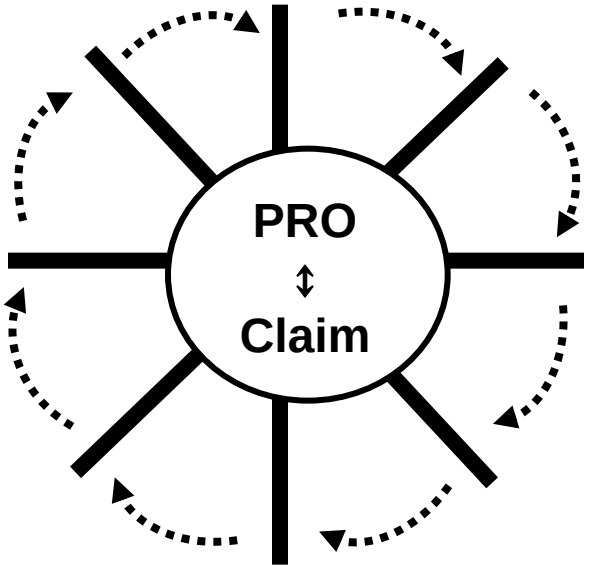
i. Hypothesize C eptual Framework

ii. Adju t Conceptual Fram rk & Draft Instru nt

iii. Confirm Conceptual Fram & Assess Othe surement Properties

iv. Col t nalyze, & In ret Data

v. Modify I ument




CAP-Questionnaire

# Symptom Assessment in CABP – Content Validity

**Literature**
- Cough
- Sputum production (color)
- Dyspnea
- Pleuritic chest pain
- Headache
- Chills
- Shaking
- Excessive sweating
- Clammy skin
- Myalgia/muscle pain
- Fatigue
- Tired
- Weak
- Nauea
- Vomiting

**Symptom Severity***

**PSS***

**Symptom Severity****
- ✓ Cough
- ✓ Dyspnea
- ✓ Sputum
- ✓ Pleuritic chest pain
- ✓ Fatigue

* Marrie et al., 2004
** Metlay et al, 1997

**CAP Questionnaire***
- ✓ Cough
- ✓ Sputum production
- ✓ Sputum color
- ✓ Sputum with ease
- ✓ Shortness of breath
- ✓ Severity of shortness of breath
- ✓ Feeling fit
- ✓ General health

* el Moussaoui et al., 2004

**CAP-SYM 12***
- ✓ Coughing
- ✓ Shortness of breath
- ✓ Chest pains
- ✓ Headache
- ✓ Chills
- ✓ Sweating
- ✓ Muscle pain
- ✓ Fatigue
- ✓ Nausea
- ✓ Lack of appetite
- ✓ Trouble concentrating
- ✓ Trouble sleeping

**CAP-SYM 18***
- ✓ Coughing up phlegm
- ✓ Vomiting
- ✓ Coughing up blood
- ✓ Diarrhea
- ✓ Stomach pain
- ✓ Trouble thinking

*Lamping et al., 2002

# CAP-Sym Development – Content Validity

- Qualitative Interviews
  - Telephone or face-to-face
  - Daily life with CAP
  - Symptoms
  - Circumstances most bothered/limited due to CAP (pre-defined format)
- N=33 with CAP
  - US & France
  - Different stages of CAP (0 – 7 days; 8-21 days; > 28 days and end of oral treatment
  - Mean age: 52 Years
  - Gender: 58% men
  - Treatment: Oral antibiotics; n=8 additional IV treatment
- Translations
  - 12 languages using forward/backward methodology



(Lamping et al., Chest 2002; 122: 920-929)

# CAP-Sym Structure

**Item Content**

**In the past 24 hours, how much have you been bothered by….**

- Coughing
- Shortness of breath
- Chest pains
- Headache
- Chills
- Sweating
- Muscle pain
- Fatigue
- Lack of appetite
- Nausea
- Trouble concentrating
- Trouble sleeping
- Coughing up phlegm
- Coughing up blood
- Vomiting
- Diarrhea
- Stomach pain
- Trouble thinking

**Response Options**

0. Did not have
1. Not at all
2. A little
3. Moderately
4. Quite a Bit
5. Extremely

**Scoring**

- Summation
- 0 to 90
- Higher Scores = Poorer Outcome

**Interviewer Administered**

(Lamping et al., Chest 2002; 122: 920-929)

# CAP-Sym Performance Properties

- N=556 CAP
  - Gender: 58% male
  - Mean Age: 50.41 (18.65) years  Range: 17-97
- Outpatient clinics, general practice, hospital centers
  - Inclusion criteria:  Fever, elevated WBC, signor or symptoms of pneumonia, and a new or progressive infiltrate on chest radiograph
- 64 Centers; 13 Countries
- RCT
  - Moxifloxacin (400 mg QD) vs Standard Treatment
  - Standard Treatment:
    - » Amoxicillin, 1g tid, and/or
    - » Clarithromycin, 500 mg bid
  - Treatment up to 14 days



(Lamping et al., Chest 2002; 122: 920-929)

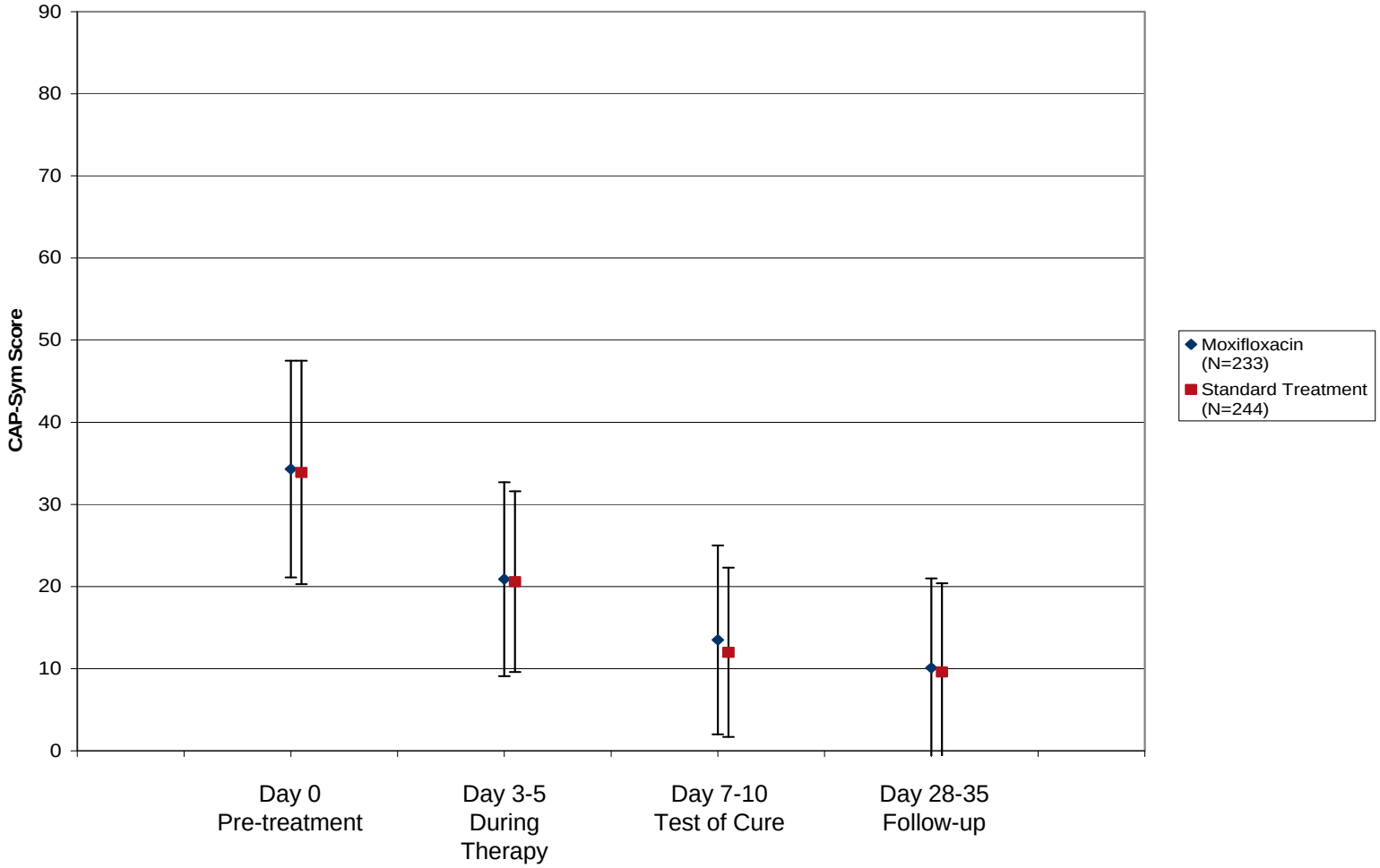
# CAP-Sym Performance Properties (N=556)

- Reliability
  - Internal Consistency — Cronbach's Alpha: 0.82
  - Test-Re-test — ICC: 0.96
- Validity
  - Content Validity — Qualitative Research
  - Construct Validity
    - » Within Scale Analyses — Alpha=0.82; Inter-item Correlations; EFA
    - » Known-Group Differences — Scores Cure > Failure (n=7) (p=0.034)
    - » Convergent Validity — Corr with SF-36 Vitality = 0.33; PCS -0.35; MSC-0.25
    - » Discriminant Validity — No corr with Age, Gender
  - Responsiveness — Change from Baseline to days 3-5; 7-10; 28-35 ES$\geq$ 1.0



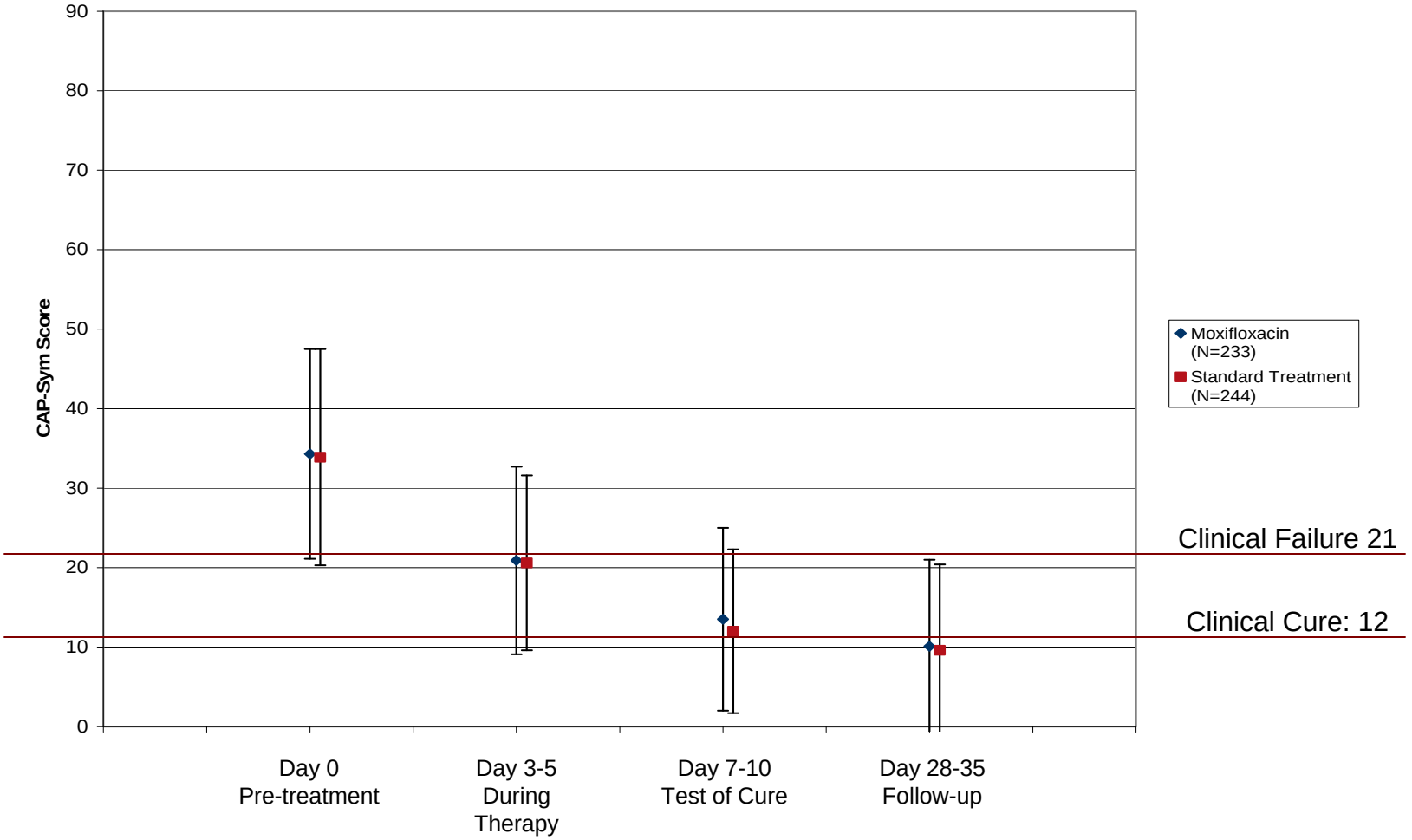
(Lamping et al., Chest 2002; 122: 920-929)

# CAP-Sym Responsiveness in the RCT




(Torres et al., Eur Respir J 2003; 21: 135-143)

# CAP-Sym Change Over Time (RCT)






(Torres et al., Eur Respir J 2003; 21: 135-143)

# The Development & Validation Process: Modified Wheel and Spokes (Simplified)

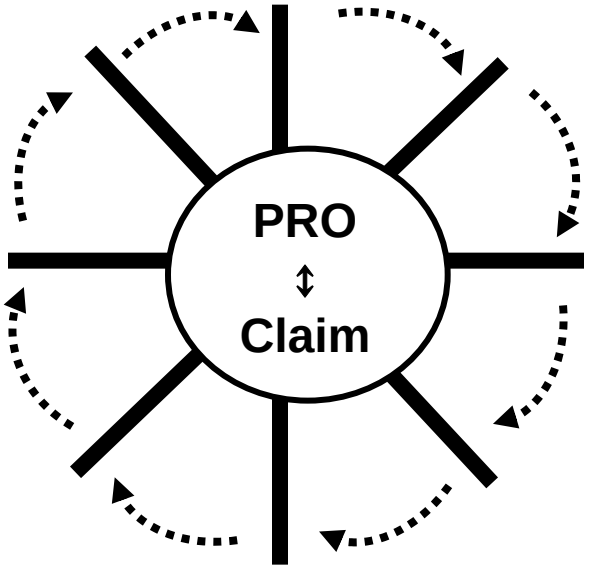
i. Hypothesize Conceptual Framework

ii. Adjust Conceptual Framework & Draft Instrument

iii. Confirm Conceptual Framework & Assess Other Measurement Properties

iv. Collect, Analyze, & Interpret Data

v. Modify Instrument




CAP-Sym

# The Development & Validation Process (CAP-Sym)

i. Hypothesize Conceptual Framework
- Outline hypothesized concepts & potential claims
- Determine the intended population
- Determine the intended application/characteristics (type of scores, mode, frequency of administration)
- Perform literature/expert review
- Develop hypothesized conceptual framework
- Position in preliminary endpoint model
- Document preliminary instrument development

iv. Collect, Analyze, & Interpret Data
- Document interpretation of treatment benefit in relation to claim

v. Modify Instrument
- Change wording of items, populations, response options, recall period, or mode/method of administration/data collection
- Translate & culturally adapt to other languages
- Evaluate modifications as appropriate
- Document all changes

# Measuring Symptoms of CABP

- Are there existing CABP Symptom PRO Instruments?
  - Yes – Examples: PSS, CAP Questionnaire, CAP-Sym
  - All responsive to change
- Can these CABP Symptom PRO Instruments be used in clinical trials evaluating the safety and efficacy of anti-infective agents?
  - Do they follow current standards for endpoint development and validation?
  - Are they suitable for clinical trials in a regulatory context?
  - CAP-Sym – Closest to FDA Draft Guidance for PRO measures
    - » Qualitative empirical foundation; quantitative evaluation
    - » Key issues – Content validity relative to target claim (“bothersome” vs “severity” rating); documentation (evaluation limited to the publication); limited information on interpretation
- What are the options?
  - Further examination of the CAP-Sym for consistency with standards
    - » Full evidence dossier for detailed assessment and regulatory review
    - » If consistent, move forward with the measure
  - Consider adapting the instrument
    - » Make adjustments and validate the modified instrument; documentation relative to guidance
  - Develop a new measure
    - » Using current standards and guidance documents

# Considerations: Population, Claims, Positioning

- Population – CABP
  - Hospitalized vs Outpatient
  - Presenting vs enriched (PORT)
  - Diagnostic criteria – Signs & Symptoms + chest radiograph?
    - » Symptoms – standardized (CABP Symptom PRO Instrument)
- Claims
  - Clinical response – Recovery
    - » Time to clinical response
    - » Clinical response (success/failure) at Day X
  - Measurement of “Recovery”
    - » Symptom resolution
    - » Sign and symptom resolution – composite symptoms + sign (e.g., afebrile)
- Positioning
  - Primary
  - Secondary

# Symptom Outcomes: Next Steps

- Options
  - Further examination of the CAP-Sym for consistency with standards
  - Consider adapting the instrument
  - Develop a new measure

- Next Steps
  - Determine the population & claim
  - Select from the options
  - Transition to the outcome
    - » Exploratory ⇨ Secondary ⇨ Primary
  - Refine or replicate during the transition

- Possible Path
  - Collaboration
  - Collaboration with the FDA through the C-Path Institute PRO Consortium
    - » http://www.c-path.org/PRO.cfm

# Clinical Response in CABP Trials

**Key questions:**

- How is “clinical response” standardized for endpoint measurement?
- How are patient-reported symptoms and clinician-observed signs standardized and quantified to determine “clinical response” to treatment in randomized, controlled trials of CABP treatment in a regulatory context?
- How *should* clinical response be standardized for endpoint measurement in multinational trials?

**Addressed:**

- Patient Reported: Symptoms of pneumonia

**To be considered:**

- Clinician Observed:  Signs of pneumonia

# Clinician-Observed: Signs of Pneumonia

- Clinician-reported outcome (ClinRO) – A standardized rating of directly observed aspects of a patient's health status that require clinical assessment and judgment.
  - Behaviors, signs, or observable symptoms
  - Definition:  Measuring Study Endpoints in Clinical Trials, DIA, New Orleans LA 2009
  - "To be meaningful, however, there should be evidence that the [PRO] instrument effectively measures the particular concept that is studied."  (US FDA PRO Draft Guidance, 2006)

# Scientific Principles

- Properties of Study Endpoints
  - Reliability – Precision
    - All elements of a given measure correspond/correlate with one another
    - Scores are stable over time in stable patients
    - Scores are reproducible across raters/observers
  - Validity – Measures what it purports to measure
    - Content Validity – Qualitative
      - How well the instrument measures the *target concept*
        - Contains the relevant & important aspects of the concept
        - "What" drives "How"
      - Evaluation – Based on the process used to develop and select items
        - Confidence in the rigor of the development methodology
    - Construct Validity – Quantitative
      - How well scores on the instrument measure (quantify) what is intended
      - Relationship to other outcome measures – similar and dissimilar
        - Known-groups; convergent, discriminant
    - Responsiveness
      - Sensitivity to change

# Reproducibility of Chest Findings

Table 1.—Precision of Physical Examination Findings in Examination of the Chest*

| Physical Examination Finding | Agreement, %† | K Value |
|---|---|---|
| Tachypnea | 63 | 0.25 |
| Reduced chest movement | 70 | 0.38 |
| Increased tactile fremitus | 85 | 0.01 |
| Dullness to percussion | 77 | 0.52 |
| Decreased breath sounds | . . ‡ | 0.43 |
| Wheezes | 79 | 0.51 |
| Crackles | 72 | 0.41 |
| Bronchial breath sounds | . . ‡ | 0.32 |
| Whispered pectoriloquy | . . ‡ | 0.11 |

Accounts for chance agreement

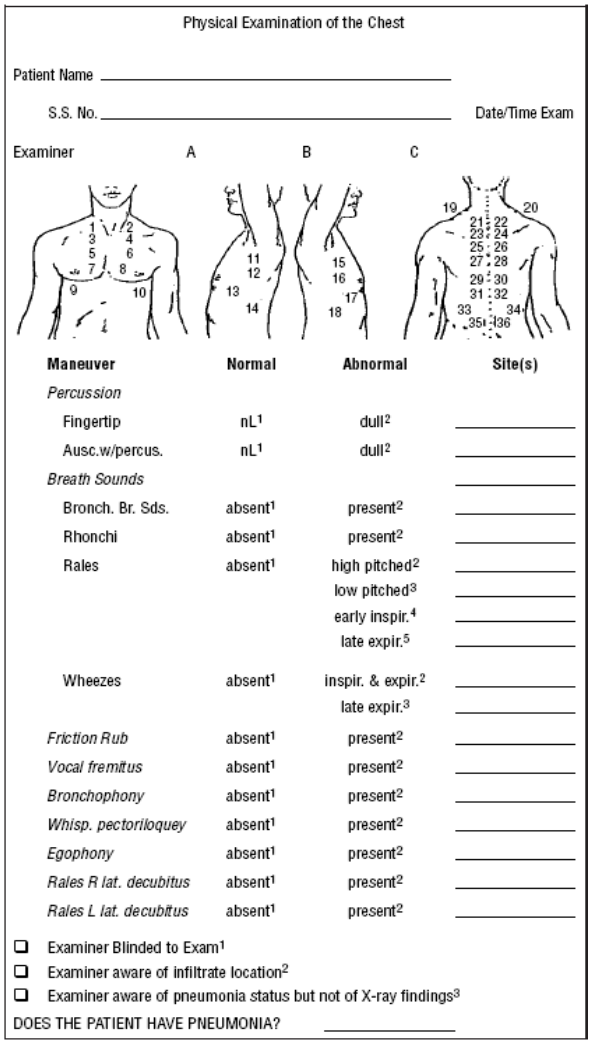
0=chance; 1=perfect agreement

*Adapted from Spiteri et al.[23]
†Calculated based on data provided in table 1 of Spirteri et al. [23]
‡ mean pair agreement rates were not calculated for the signs for which 2 or more physicians in a group failed to report the presence or absence of the sign.



(Spirteri et al., Lancet 1988, in Metlay et al, JAMA 1997; 1440-1445)

# Reproducibility of Chest Findings - Pneumonia



Physical Examination of the Chest

Patient Name _____

S.S. No. _____ Date/Time Exam

Examiner A B C

| Maneuver | Normal | Abnormal | Site(s) |
|---|---|---|---|
| *Percussion* | | | |
| Fingertip | nL[1] | dull[2] | |
| Ausc.w/percus. | nL[1] | dull[2] | |
| *Breath Sounds* | | | |
| Bronch. Br. Sds. | absent[1] | present[2] | |
| Rhonchi | absent[1] | present[2] | |
| Rales | absent[1] | high pitched[2] | |
| | | low pitched[3] | |
| | | early inspir.[4] | |
| | | late expir.[5] | |
| Wheezes | absent[1] | inspir. & expir.[2] | |
| | | late expir.[3] | |
| *Friction Rub* | absent[1] | present[2] | |
| *Vocal fremitus* | absent[1] | present[2] | |
| *Bronchophony* | absent[1] | present[2] | |
| *Whisp. pectoriloquey* | absent[1] | present[2] | |
| *Egophony* | absent[1] | present[2] | |
| *Rales R lat. decubitus* | absent[1] | present[2] | |
| *Rales L lat. decubitus* | absent[1] | present[2] | |

☐ Examiner Blinded to Exam[1]
☐ Examiner aware of infiltrate location[2]
☐ Examiner aware of pneumonia status but not of X-ray findings[3]

DOES THE PATIENT HAVE PNEUMONIA? _____

*Data sheet showing sites of chest physical examination and recorded maneuvers performed at each site. The site of any physical finding was identified by the numbers in the diagrams.*

**Table 3. Physician Agreement on Findings as Reflected by κ Values**

| | Physician A vs Physician B | Physician B vs Physician C | Physician A vs Physician C |
|---|---|---|---|
| Bronchial breath sounds | | | |
| Left lung | 0.14 | −0.14 | −0.14 |
| Right lung | 0.03 | 0.07 | 0.14 |
| Bronchophony | | | |
| Left lung | −0.12 | −0.06 | −0.14 |
| Right lung | 0.16 | 0.25 | 0.22 |
| Egophony | | | |
| Left lung | −0.08 | 0 | 0.18 |
| Right lung | 0.03 | −0.10 | 0.03 |
| Rales | | | |
| Left lung | 0.35 | 0.64 | 0.51 |
| Right lung | 0.24 | 0.49 | 0.65 |
| Lateral decubitus rales | | | |
| Left lung | 0.23 | 0.47 | 0.23 |
| Right lung | 0.32 | 0.39 | 0.39 |
| Wheezes | | | |
| Left lung | 0.49 | 1.0 | 0.65 |
| Right lung | 0.17 | 0.65 | −0.05 |
| Rhonchi | | | |
| Left lung | 0.13 | 0 | −0.06 |
| Right lung | 0.18 | −0.05 | −0.05 |
| Percussion (fingertip) | | | |
| Left lung | 0 | * | 0 |
| Right lung | 0.40 | 1.0 | 0.28 |
| Percussion (auscultatory) | | | |
| Left lung | 0 | 0 | 0.45 |
| Right lung | −0.04 | 0.28 | 0.10 |
| Pneumonia diagnosis (% agreement) | 0.18 (60) | 0.31 (69) | 0.43 (72) |

*$_{\kappa}$ *is not defined due to expected agreement of 1.0.*

K: 0=chance;
1=perfect agreement

(Wipf et al, Arch Intern Med 1999, 1082-1087)

# Clinician-Observed: Signs of Pneumonia
# A Road Map for Standardization?

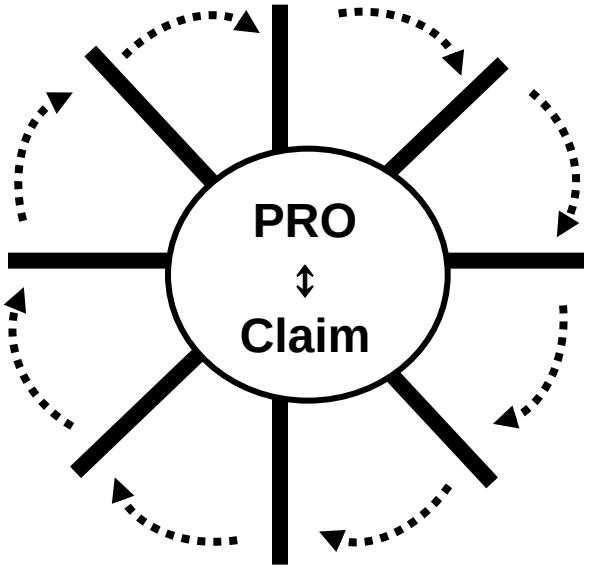
i. Hypothesize Conceptual Framework

ii. Adjust Conceptual Framework & Draft Instrument

iii. Confirm Conceptual Framework & Assess Other Measurement Properties

iv. Collect, Analyze, & Interpret Data

v. Modify Instrument



# Standardizing Clinical Response in CABP: Key Questions

- What is the Population?
  - Hospitalized vs Outpatient
  - Presenting vs enriched (PORT)
  - Diagnostic criteria – Signs & Symptoms + chest radiograph?
    - » Symptoms – standardized (CABP Symptom PRO Instrument)
- What are the Claims?
  - Clinical response – Recovery
    - » Time to clinical response
    - » Clinical response (success/failure) at Day X
  - Measurement of “Recovery”
    - » Symptom resolution
    - » Sign and symptom resolution – composite symptoms + sign (e.g., afebrile)
- How are the Outcomes Positioned?
  - Primary/secondary?  In the short and long term?

# Overview/Summary

- Introductory Comments
  - Scientific Principles - Efficacy; CABP; Health Outcomes, Properties of Study Endpoints
  - Health Outcomes/Endpoints in CABP
  - Properties of Study Endpoints – PROs
  - Clinical Response in CABP
- PRO Instruments & Development/Regulatory Context
  - Development and validation of a PRO
  - Measuring symptoms of CABP – PROs
- Existing PRO Instruments for CABP
  - Pneumonia Symptom Severity Scales (Metlay et al, 1997; Marrie et al., 2004)
  - The Community-Acquired Pneumonia (CAP) questionnaire (el Moussaoui et al., 2004)
  - The Community-Acquired Pneumonia Symptom Questionnaire (Lamping et al., 2002)
  - Next Steps
- Clinical Response in CABP Trials
  - Clinician-Observed Outcomes
  - Standardization – Key Questions

# Conclusions


- John Glenn, Friendship 7
- February 20, 1962, Cape Canaveral
- First American to Orbit Earth

- Space Shuttle Endeavour
- June 15, 2002

# References (1 of 2)


- Dean NC. et al. Comparing gatifloxacin an clarithromycin in pneumonia symptom resolution and process of care. Antimicrobial Agents & Chemotherapy 2006; 50:1164-1169.
- El Moussaoui R, et al. Long-term symptom recovery and health-related quality of life in patients with mild-to-moderate-severe community-acquired pneumonia. Chest 2006; 130:1165-1172.
- El Moussaoui R et al. Development & validation of a short questionnaire in community-acquired pneumonia. Thorax 2004; 59:591-595.
- Lamping et al. The Community-Acquired Pneumonia Symptom Questionnaire, Chest 2002; 122:920-929.
- Mandell LA. Infectious Diseases Society of America/American Thoracic Society Consensus Guidelines on the Management of Community-Acquired Pneumonia in Adults. Clinical Infectious Disease 2007; 44:S27-72.
- Marrie TJ et al. Resolution of symptoms in patients with community-acquired pneumonia treated on an ambulatory basis. Journal of Infection 2004; 49:302-309.
- Metlay JP et al. Measuring symptomatic and functional recovery in patients with community-acquired pneumonia. J Gen Intern Med 1997; 12:423-430.
- Metlay JP et al. Does this patient have community-acquired pneumonia? JAMA 1997; 278:1140-1445.

# References (2 of 2)


- Measuring Study Endpoints in Multinational Clinical Trials: Outcomes Reported from the Viewpoint of the Clinician, Patient, and Caregiver. Drug Information Association (DIA). New Orleans, LA, October 2009.
- Patrick D et al. Patient-reported outcomes to support medical product labeling claims: FDA perspective. Value in Health 2008; 10 (Supp 2): S125-S137.
- Rothman M et al. Use of existing PRO instruments and their modification: ISPOR Task Force Report. Value in Health 2009 [Epub ahead of print].
- Torres A et al. Effectiveness of oral moxifloxacin in standard first-line therapy in community-acquired pneumonia. Eur Respir J 2003; 21:135-143.
- US Food and Drug Administration. Guidance for Industry: Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims. http://www.fda.gov/ScienceResearch/SpecialTopics/RunningClinicalTrials/ProposedRegulationsandDraftGuidances/default.htm .
- Wipf JE et al. Diagnosing pneumonia by physical examination: Relevant or relic, Arch Intern Med 1999; 159:1082-1087.